

## ANALISIS KEPUASAN PELANGGAN MELALUI SENTIMEN ULASAN MENGUNAKAN ALGORITMA NAIVE BAYES

**Muhamad Arsadhana<sup>1</sup>, Bahtiar Efendi<sup>2</sup>, M. Trihudyatmanto<sup>3</sup>**

<sup>1,2,3</sup> Fakultas Ekonomi dan Bisnis, Universitas Sains Al-Qur'an

email<sup>1</sup>: [arsadhanamuhammad@gmail.com](mailto:arsadhanamuhammad@gmail.com)

email<sup>2</sup>: [bahtiarefd@unsiq.ac.id](mailto:bahtiarefd@unsiq.ac.id)

email<sup>3</sup>: [trihudyatmanto@unsiq.ac.id](mailto:trihudyatmanto@unsiq.ac.id)

### ABSTRACT

The advancement of digital technology has led to a large amount of data being generated by active internet users. In this study, the focus is on analysing customer sentiment towards MSMEs selling on the Tokopedia platform. The analysis is conducted using the naïve bayes algorithm method and Net Promoter Score (NPS) to measure customer satisfaction. The findings reveal that 75.69% of customer sentiment is positive, 11.08% is neutral, and 13.23% is negative. The Multinomial Naive Bayes model utilized achieves an accuracy rate of 80%, with precision, recall, and F1-Score all reaching 80%. Moreover, the calculated NPS value is 62.46%, indicating that a majority of customers are satisfied with the MSMEs on Tokopedia

*Keywords: Sentiment Analysis, Naive Bayes, MSME, Tokopedia, Customer Satisfaction*

### ABSTRAK

Kemajuan teknologi digital telah menghasilkan sejumlah besar data yang dihasilkan oleh pengguna internet aktif. Pada penelitian ini, fokusnya adalah menganalisis sentimen pelanggan terhadap UMKM yang berjualan di platform Tokopedia. Analisis dilakukan dengan menggunakan metode algoritma naïve bayes dan Net Promoter Score (NPS) untuk mengukur kepuasan pelanggan. Temuan menunjukkan bahwa 75,69% dari sentimen pelanggan adalah positif, 11,08% netral, dan 13,23% negatif. Model Multinomial Naive Bayes yang digunakan mencapai tingkat akurasi 80%, dengan presisi, recall, dan F1-Score yang mencapai 80%. Selain itu, nilai NPS yang dihitung adalah 62,46%, menunjukkan bahwa mayoritas pelanggan puas dengan UMKM di Tokopedia

Kata kunci: Analisis Sentimen, Naive Bayes, UMKM, Tokopedia, Kepuasan Pelanggan

### 1. Pendahuluan

Berkembangnya teknologi digital menjadikan banyaknya pengguna aktif internet pada saat ini membuat jumlah data yang dihasilkan sangat besar (big data) hal ini menciptakan peluang bagi para pengambil keputusan untuk memperoleh wawasan lebih baik (Miah et al., 2017). Jumlah data yang sangat besar dihasilkan oleh teknologi yang berkembang secara terus-menerus. Apabila data tersebut dapat diproses dan digunakan, data tersebut dapat memberikan informasi yang bermanfaat (Fan et al., 2014). Analisis Big Data memungkinkan UMKM untuk memperoleh pemahaman yang lebih mendalam tentang perilaku pelanggan, tren pasar, dan strategi bisnis yang lebih baik (Bhatti et al., 2022) Dalam beberapa waktu terakhir, inovasi teknologi open source, perangkat keras komoditas, komputasi awan, aplikasi berbasis data, teknik visualisasi telah mendorong transformasi dalam model

bisnis dan ruang teknologi data (Jha & Sahoo, 2021).

Dalam era transformasi digital, peran E-commerce bagi UMKM menjadi semakin penting sebagai strategi pemasaran digital (Sugiharto, 2024). Selama transisi dari bisnis konvensional ke digital, perusahaan harus mempertimbangkan rencana strategis untuk memastikan bahwa proses bisnis baru yang ada tetap berkelanjutan (Murdiana & Hajaoui, 2020). E-commerce adalah platform pemasaran yang memanfaatkan situs web untuk melakukan transaksi atau mendukung proses penjualan produk secara online (Kotler Philip & Keller Kevin Lane, 2016). Banyak platform e-commerce menampilkan produk dan layanan dengan gambar dan review pengguna (Jain Vipin et al., 2021). E-commerce berada di garis depan dalam mengubah strategi pemasaran, berdasarkan teknologi baru, dan memfasilitasi informasi produk dan pengambilan keputusan yang lebih baik.

Tokopedia adalah salah satu platform e-commerce terkenal di Indonesia yang menjual berbagai produk dan jasa (Yusuf et al., 2022). Tokopedia lebih berfokus pada pengembangan pendekatan yang lebih berpusat pada konsumen dari pada promosi (Hirzy & Dellyana, 2023). Platform e-commerce seperti Tokopedia telah menjadi solusi yang menarik bagi para pengusaha yang berupaya menghadapi tantangan dan memanfaatkan peluang baru di era digital (Pangilinan et al., 2023) Untuk bertahan dalam industri e-commerce saat ini, bisnis online harus memahami elemen penting apa saja yang menentukan kepuasan pelanggan (Masyhuri, 2022)

Kepuasan pelanggan dapat diukur melalui ulasan atau review yang ditinggalkan oleh pengguna di platform e-commerce (Alqurafi & Alsanoosy, 2024). Kepuasan pelanggan, istilah yang sering digunakan dalam pemasaran, merujuk pada ukuran bagaimana produk atau layanan yang disediakan oleh perusahaan memenuhi atau melampaui harapan pelanggan (Haralayya, 2021). Kepuasan pelanggan penting bagi sebuah perusahaan karena pengaruhnya terhadap kondisi keuangan perusahaan dan dampak positifnya terhadap persepsi publik terhadap perusahaan (Basari & Shamsudin, 2020) Perusahaan diharuskan untuk mendapatkan pemahaman yang lebih baik tentang hubungan lingkungan online antara kepuasan dan niat perilaku, dan untuk menetapkan strategi pemasaran online antara inisiatif kepuasan dan program niat perilaku (Agarwal & Gowda, 2021).

Penelitian ini bertujuan untuk menganalisis kepuasan pelanggan melalui analisis sentiment, analisis sentimen adalah proses mengumpulkan dan menganalisis pendapat, pemikiran, dan kesan orang mengenai berbagai topik, produk, subjek, dan layanan (Wankhade et al., 2022). Analisis sentimen, juga disebut Opinion Mining, adalah pekerjaan untuk mengekstraksi dan menganalisis pendapat, sentimen, sikap, dan persepsi individu terhadap berbagai hal seperti topik, barang, dan layanan (Birjali et al., 2021). Analisis sentimen digunakan dalam berbagai

pendekatan seperti ulasan produk dan layanan (Abualigah et al., 2020).

Algoritma Naïve Bayes dipilih sebagai model untuk analisis sentiment, Naive Bayes adalah algoritma klasifikasi untuk masalah klasifikasi multikelas, disebut Naive Bayes karena perhitungan probabilitas untuk setiap kelas disederhanakan agar perhitungannya mudah dilakukan (Viet et al., 2021). Metode Naive Bayes merupakan salah satu metode yang cukup populer untuk klasifikasi teks karena asumsi kisi-kisi yang efektif, implementasi yang cepat dan mudah (Abbas et al., 2019). Hasil empiris menunjukkan bahwa selective naïve Bayes menunjukkan akurasi klasifikasi yang unggul, namun pada saat yang sama mempertahankan kesederhanaan dan efisiensi (Chen et al., 2020).

Pengujian kepuasan pelanggan melalui sentiment ulasan menggunakan metode Naïve Bayes dilakukan menggunakan indikator Net Promoter Score, Reichheld mempresentasikan NPS dalam artikel Harvard Business Review berjudul “The Only Number You Need to Grow” (Reichheld, 2003) Net Promoter Score (NPS) adalah konsep yang semakin populer karena kesederhanaannya dan klaimnya tentang hubungan dengan profitabilitas. Konsep ini telah terbukti oleh hasil replikasi independen oleh peneliti lain, serta adopsi oleh perusahaan-perusahaan terkemuka seperti GE, American Express, dan Microsoft (Creamer, 2006). Salah satu kekuatan lain dari NPS adalah kesesuaiannya dengan era komunikasi konsumen-ke-konsumen dan media yang dihasilkan oleh konsumen seperti blog dan pemasaran viral (Kirby & Marsden, 2006).

Penelitian terdahulu yang berjudul “Naives Bayes Algorithm for Twitter Sentiment Analysis” menghasilkan bahwa sentimen negatif lebih dominan pada periode tersebut sebesar 46%, sementara sentimen positif lebih dominan sebesar 35% dan sentimen netral sebesar 20% (Samsir et al., 2021). Penelitian yang berjudul “Sentiment Analysis On Shopee E-Commerce Using The Naïve Bayes Classifier Algorithm ” Hasil akurasi

klasifikasi menggunakan pendekatan Naive Bayes Classifier berkisar antara 88,17% hingga 99,76% (Saputri & Februariyanti, 2022). Sedangkan pada penelitian yang berjudul “Improvement Of Naive Bayes Algorithm In Sentiment Analysis Of Shopee Application Reviews On Google Play Store“ Hasil pengujian menunjukkan nilai akurasi sebesar 81%, hasil ini menunjukkan bahwa aplikasi shopee ulasan yang diberikan cenderung negative (Elistiana et al., 2023).

Berdasarkan latar belakang tersebut bahwa pelanggan sering kali mengekspresikan kepuasan atau ketidakpuasan mereka melalui ulasan yang ditinggalkan di platform Tokopedia. Ulasan ini mengandung sentimen yang dapat mempengaruhi kepuasan pelanggan terhadap UMKM. Oleh karena itu, muncul fenomena penting yang perlu diteliti, yaitu bagaimana sentimen pelanggan terhadap kepuasan pelanggan. Hasil penelitian ini juga diharapkan dapat digunakan oleh UMKM untuk memahami kepeuasan pelanggan sehingga mampu meningkatkan keberlanjutan bisnis mereka di era digital.

## 2. Tinjauan Pustaka

### Analisis Sentiment

Sentiment Analisis atau yang bisa disebut Opinion Mining adalah studi komputasi tentang opini, sikap, dan emosi orang terhadap suatu entitas (Medhat et al., 2014). SA, yang juga dikenal sebagai analisis opini atau penambangan opini, merupakan bidang penting dalam Natural Language Processing (NLP). Teknologi ini dirancang untuk secara otomatis mengekstrak dan menganalisis sentimen serta pandangan yang terdapat dalam teks (Chaturvedi et al., 2018). Pada saat yang sama, sudah menjadi praktik umum untuk menggunakan jejaring sosial untuk mengekspresikan pendapat seseorang tentang topik atau peristiwa yang menarik, serta memberikan umpan balik tentang produk dan layanan yang digunakan (Luneva et al., 2015).

### Customer Satisfaction

Konsep kepuasan pelanggan ditinjau dari perspektif terapan berbasis teori yang didasarkan pada sejumlah besar karya dalam komunitas akademis yang baru-baru ini menemukan jalannya ke dalam survei non akademis (Oliver, 2010). Kepuasan pelanggan dapat didefinisikan sebagai evaluasi keseluruhan yang didasarkan pada pengalaman konsumen saat membeli barang atau jasa dan mengkonsumsinya secara keseluruhan dari waktu ke waktu (Fornell et al., 1996). Dalam pemasaran, kepuasan pelanggan juga menjadi bagian dari proses tersebut, yang berarti memastikan harapan pelanggan terhadap bagaimana barang dan jasa difasilitasi oleh perusahaan. Oleh karena itu, informasi yang dapat ditindak lanjuti tentang cara membuat pelanggan lebih puas menjadi hasil yang sangat penting (Oliver, 1999).

## 3. Metode Penelitian

### Pendekatan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif deskriptif yang dilakukan untuk menggambarkan suatu hasil penelitian (Ramdhan, 2021) Dengan metode data mining untuk menganalisis sentimen ulasan pelanggan pada UMKM Oleh-Oleh Wonosobo yang berada di platform Tokopedia. Data mining adalah alat dan algoritma yang digunakan untuk memungkinkan analisis dan pengguna akhir untuk memecahkan masalah yang jika tidak, akan membutuhkan banyak upaya manual atau tidak akan terpecahkan (Mining, 2006). Analisis dilakukan menggunakan algoritma Naive Bayes untuk mengklasifikasikan sentimen menjadi positif dan negative. Algoritma ini menggunakan teorema Bayes, sebuah formula yang menghitung probabilitas dengan menghitung frekuensi nilai dan kombinasi nilai dalam data historis (Viet et al., 2021b). perhitungan nilai dapat dilihat dalam rumus sebagai berikut :

$$P(C | X) = \frac{P(X|C) \cdot P(C)}{P(X)}$$

**P(C|X):** Probabilitas bahwa data X termasuk dalam kelas C (posterior probability).

**P(X|C)**): Probabilitas data X muncul, dengan asumsi bahwa data tersebut berasal dari kelas C (likelihood).

**P(C)** Probabilitas awal kelas C (prior probability).

**P(X)** Probabilitas data X muncul secara keseluruhan (evidence).

### Sumber data

Data ulasan pelanggan diambil dari Toko Oleh Oleh Wonosobo yang berada di Tokopedia

### Metode Pengumpulan Data

Data diambil menggunakan web scraping, web scraper berfungsi mengambil tautan dari web, kemudian data diekstrak untuk mendapatkan data dari tautan sumber dan akhirnya menyimpan data tersebut ke dalam file csv(Thomas & Mathur, 2019). Data hasil scraping berisi 469 ulasan pelanggan dari toko Oleh-Oleh Wonosobo di Tokopedia.

### Alat Analisis

Pada penelitian menggunakan alat analisis Bahasa pemrograman Phyton, Python adalah bahasa pemrograman yang mudah dipelajari dan dipahami, bahkan bagi pemula, sintaksnya yang sederhana membuatnya sangat populer, selain itu, Python juga sangat fleksibel dan bisa digunakan untuk berbagai macam proyek (Python, 2021).

### Preprocessing Data

Rangkaian teknik yang digunakan sebelum penerapan metode data mining dinamakan sebagai data preprocessing untuk data mining (García et al., 2015) Langkah awal penting dalam analisis data adalah preprocessing data, yang mencakup berbagai teknik untuk membersihkan, mengubah, dan menyusun ulang data. Agar properti yang diminati dapat diprediksi dengan tepat, pemrosesan awal diperlukan (Mishra et al., 2020; Tri Jaka, 2015). Setelah preprocessing, data ulasan pelanggan telah diolah menjadi lebih bersih dan siap untuk analisis lebih lanjut. Total data setelah preprocessing adalah 325 ulasan. Ini menunjukkan bahwa dari 469 data awal, terdapat pengurangan kemungkinan akibat penghapusan data duplikat, ulasan kosong, atau ulasan yang tidak relevan.

Terjemahan ke bahasa Inggris telah menormalkan frasa informal ke bahasa yang lebih standar.

### Analisis Sentimen

Labeling Data: Data akan diberi label secara otomatis berdasarkan ulasan sentiment positif dan negative. Metode Klasifikasi: Menggunakan Naive Bayes Algorithm yang terkenal untuk analisis sentimen karena sederhana dan efisien untuk teks pendek.

### Evaluasi Model

Membagi data menjadi dua bagian Training Set dan Data Uji selanjutnya Menggunakan metrik evaluasi seperti: Akurasi, Precision, Recall, F1-Score

### Implementasi Hasil

Untuk mengukur dan memetakan loyalitas pelanggan menggunakan metode NPS. Metode untuk menghitung indeks NPS sangat sederhana dan mudah untuk diinterpretasikan. Ini dihitung dengan mengurangi detractor dari promoter—passive tidak diperhitungkan

$$\text{NPS (\%)} = \frac{\text{Promoters} - \text{Detractros}}{\text{The numbers of respondents}} \times 100$$

Indeks persentase ini akan didefinisikan sebagai berikut:

**Promoters = jumlah sentimen positif**

**Detractros = Jumlah respon negatif**

**The numbers of respondents = Total seluruh jumlah data ulasan**

Pada penelitian ini sentimen netral tidak diguakan karena fokus utama yang digunakan pada sentiment yang memiliki dampak nyata bagi kepuasan . Oleh karena itu NPS, nilainya bisa positif maupun negatif. Interpretasinya sama mudahnya untuk dianalisis (Grisaffe, 2007).

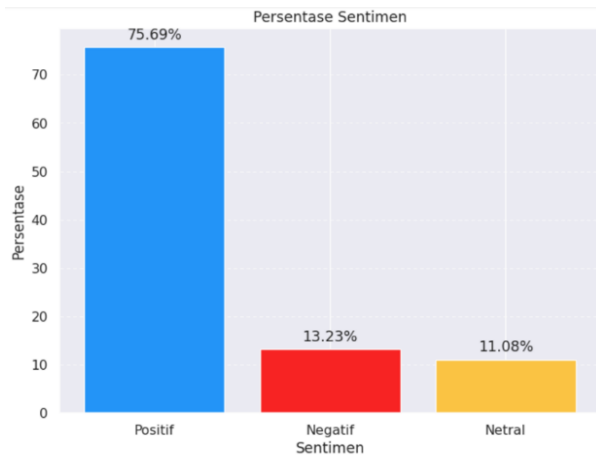
## 4.Hasil dan Pembahasan

### Analisis Sentimen

Dari 325 ulasan pelanggan di Toko Oleh-Oleh Wonosobo di Tokopedia, terdapat 246

ulasan bersentimen positif, 36 ulasan bersentimen netral, dan 43 ulasan bersentimen negatif. Sebagian besar ulasan bersentimen positif menunjukkan tingkat ulasan positif yang cukup tinggi, didukung oleh hasil visualisasi dari WorldCloud yang digunakan pada bahasa pemrograman Python, seperti yang ditunjukkan pada Gambar 1.

**Gambar 1.** Grafik Hasil analisis sentimen



Sumber : Hasil penelitian di olah dengan python, 2024

Berdasarkan analisis sentiment, grafik menunjukkan persentase sentimen pelanggan. Sentimen positif 75.69% menunjukkan bahwa sebagian besar ulasan pelanggan memiliki respons positif, sentimen negatif 13.23% menunjukkan bahwa sebagian kecil ulasan memiliki respons negatif, dan sentimen netral 11.08% menunjukkan bahwa beberapa ulasan bersifat netral. Dari grafik ini, dapat disimpulkan bahwa ulasan sentimen pada penelitian ini cukup banyak yang bersifat positif.

**Evaluasi Model Algoritma Multinomial Naive Bayes**

Dalam klasifikasi sentimen, model Multinomial Naive Bayes sangat baik, terutama dengan akurasi dan recall yang tinggi. Algoritma ini digunakan untuk klasifikasi sentimen dan menghasilkan performa evaluasi dengan F1-score 71.11%, akurasi sebesar 80%, ketepatan sebesar 64%, dan recall sebesar 80%. Hasil evaluasi menunjukkan bahwa

model ini mampu mengidentifikasi sentimen positif dengan cukup baik.

**Gambar 2.** Hasil Pemodelan Naïve Bayes

model	AccuracyScore	Precision	Recall	F1Score
0 Multinomial Naive Bayes	0.8	0.64	0.8	0.711111

Sumber : Hasil penelitian di olah dengan python, 2024

Berdasarkan hasil evaluasi model, nilai Accuracy 80% menunjukkan bahwa model memiliki kemampuan untuk mengklasifikasikan ulasan dengan benar sebanyak 80% dari total data uji. Ini adalah hasil yang cukup baik untuk model berbasis teks yang memiliki dataset yang relatif kecil. Seberapa akurat model dapat memprediksi kelas tertentu, terutama untuk sentimen positif, disebut precision. Dengan nilai 0.64, sekitar 64% prediksi positif yang dihasilkan model benar-benar positif; pengecualian menunjukkan seberapa baik model mendeteksi semua data sebenarnya yang positif; dan nilai 0,8 menunjukkan bahwa model berhasil mendeteksi 80% dari semua ulasan positif yang ada. F1-Score, metrik yang menggabungkan precision dan recall, memiliki nilai 0.711, menunjukkan bahwa model memiliki keseimbangan yang baik antara kemampuan untuk menemukan ulasan positif (recall) dan memberikan prediksi yang relevan (precision).

**Implementasi Hasil**

Dalam penelitian ini, nilai Net Promoter Score (NPS) digunakan untuk mengukur kepuasan pelanggan.

$$NPS (\%) = \frac{246 - 43}{325} \times 100$$

$$NPS = 62.46\%$$

Hasil perhitungan Net Promoter Score (NPS) pada penelitian ini sebesar 62.46% menunjukkan bahwa sebagian besar pelanggan tidak hanya puas dengan toko tersebut, tetapi mereka juga akan dengan senang hati merekomendasikan toko tersebut kepada orang lain. Loyalitas pelanggan yang tinggi sangat penting untuk keberlanjutan bisnis, terutama dalam hal memperoleh pelanggan

tetap dan menarik pelanggan baru melalui rekomendasi.

## 5. Kesimpulan

Dengan menggunakan pendekatan analisis sentimen berbasis algoritma Multinomial Naive Bayes, penelitian ini berhasil menganalisis kepuasan pelanggan terhadap loyalitas pelanggan. Hasil menunjukkan tingkat sentimen positif sebesar 75,69%, tingkat sentimen netral sebesar 11,08%, dan tingkat sentimen negatif 13,23%. Model Multinomial Naive Bayes yang digunakan memiliki tingkat akurasi 80%, precision 64%, recall 80%, dan F1-Score 80% persen. nilai Net Promoter Score (NPS) yang dihitung mencapai 62.46%, yang menunjukkan bahwa mayoritas pelanggan merasa puas. Hasil analisis sentimen negatif dapat dimanfaatkan pelaku UMKM untuk merancang strategi peningkatan kepuasan pelanggan. penelitian selanjutnya dapat membandingkan kinerja algoritma Multinomial Naive Bayes dengan model lain seperti logistic regression, Random Forest, atau deep learning.

## DAFTAR PUSTAKA

- Abbas, M., Memon, K. A., Jamali, A. A., Memon, S., & Ahmed, A. (2019). Multinomial Naive Bayes classification model for sentiment analysis. *IJCSNS Int. J. Comput. Sci. Netw. Secur*, *19*(3), 62.
- Abualigah, L., Alfar, H. E., Shehab, M., & Hussein, A. M. A. (2020). Sentiment analysis in healthcare: a brief review. *Recent Advances in NLP: The Case of Arabic Language*, 129–141.
- Agarwal, I., & Gowda, K. R. (2021). The effect of airline service quality on customer satisfaction and loyalty in India. *Materials Today: Proceedings*, *37*, 1341–1348. <https://doi.org/https://doi.org/10.1016/j.matpr.2020.06.557>
- Alqurafi, A., & Alsanoosy, T. (2024). Measuring Customers' Satisfaction Using Sentiment Analysis: Model and Tool. *J. Comput. Sci*, *20*, 419–430.
- Basari, M. A. M. D., & Shamsudin, M. F. (2020). Does customer satisfaction matters? *Journal of Undergraduate Social Science and Technology*, *2*(1).
- Bhatti, S. H., Ahmed, A., Ferraris, A., Hirwani Wan Hussain, W. M., & Wamba, S. F. (2022). Big data analytics capabilities and MSME innovation and performance: A double mediation model of digital platform and network capabilities. *Annals of Operations Research*, 1–24.
- Birjali, M., Kasri, M., & Beni-Hssane, A. (2021). A comprehensive survey on sentiment analysis: Approaches, challenges and trends. *Knowledge-Based Systems*, *226*, 107134. <https://doi.org/https://doi.org/10.1016/j.knosys.2021.107134>
- Chaturvedi, I., Cambria, E., Welsch, R. E., & Herrera, F. (2018). Distinguishing between facts and opinions for sentiment analysis: Survey and challenges. *Information Fusion*, *44*, 65–77.
- Chen, S., Webb, G. I., Liu, L., & Ma, X. (2020). A novel selective naïve Bayes algorithm. *Knowledge-Based Systems*, *192*, 105361. <https://doi.org/https://doi.org/10.1016/j.knosys.2019.105361>
- Creamer, M. (2006). Do you know your score: your business rides on how consumers answer the one crucial question. *Advertising Age*, *77*(1), 24.
- Elistiana, K. M., Kusuma, B. A., Subarkah, P., & Rozaq, H. A. A. (2023). Improvement of Naive Bayes Algorithm in Sentiment Analysis of Shopee Application Reviews on Google Play Store. *Jurnal Teknik Informatika (Jutif)*, *4*(6), 1431–1436.
- Fan, J., Han, F., & Liu, H. (2014). Challenges of Big Data analysis. In *National Science Review* (Vol. 1, Issue 2, pp. 293–314). Oxford University Press. <https://doi.org/10.1093/nsr/nwt032>
- Fornell, C., Johnson, M. D., Anderson, E. W., Cha, J., & Bryant, B. E. (1996). The American Customer Satisfaction Index: Nature, Purpose, and Findings. *Journal of Marketing*, *60*(4), 7–18. <https://doi.org/10.1177/002224299606000403>
- García, S., Luengo, J., & Herrera, F. (2015). *Data preprocessing in data mining* (Vol. 72). Springer.

- Grisaffe, D. B. (2007). Questions about the ultimate question: conceptual considerations in evaluating Reichheld's net promoter score (NPS). *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 20, 36–53.
- Haralayya, B. (2021). Customer Satisfaction at M/s Sindol Bajaj Bidar. *Iconic Research And Engineering Journals*, 4(12), 157–169.
- Hirzy, H., & Dellyana, D. (2023). The Impact of Consumer Purchase Behavior towards User Purchase at E-commerce (Tokopedia). *European Journal of Business and Management Research*, 8(4), 154–162.
- Jain Vipin, Malviya Bindoo, & Arya Satyendra. (2021). An Overview of Electronic Commerce (e-Commerce). *Journal of Contemporary Issues in Business and Government*, 27(3). <https://doi.org/10.47750/cibg.2021.27.03.090>
- Jha, R. S., & Sahoo, P. R. (2021). Influence of big data capabilities in knowledge management—MSMEs. *ICT Systems and Sustainability: Proceedings of ICT4SD 2020, Volume 1*, 513–524.
- Kirby, J., & Marsden, P. (2006). *Connected marketing: the viral, buzz and word of mouth revolution*. Elsevier.
- Kotler Philip, & Keller Kevin Lane. (2016). *Marketing Management*. <http://www.pearsonmylabandmastering.com>
- Luneva, E. E., Yefremov, A. A., & Banokin, P. I. (2015). AUTOMATED SENTIMENT ANALYSIS EVALUATION OF SOCIAL NETWORK USERS BASED ON FUZZY LOGIC. *Statistics and Economics*, 3, 249–254.
- Masyhuri, M. (2022). Key drivers of customer satisfaction on the e-commerce business. *East Asian Journal of Multidisciplinary Research*, 1(4), 657–670.
- Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4), 1093–1113. <https://doi.org/https://doi.org/10.1016/j.asej.2014.04.011>
- Miah, S. J., Vu, H. Q., Gammack, J., & McGrath, M. (2017). A Big Data Analytics Method for Tourist Behaviour Analysis. *Information and Management*, 54(6), 771–785. <https://doi.org/10.1016/j.im.2016.11.011>
- Mining, W. I. D. (2006). *Introduction to data mining*. Springer.
- Mishra, P., Biancolillo, A., Roger, J. M., Marini, F., & Rutledge, D. N. (2020). New data preprocessing trends based on ensemble of multiple preprocessing techniques. *TrAC Trends in Analytical Chemistry*, 132, 116045. <https://doi.org/https://doi.org/10.1016/j.trac.2020.116045>
- Murdiana, R., & Hajaoui, Z. (2020). E-Commerce marketing strategies in industry 4.0. *International Journal of Business Ecosystem & Strategy (2687-2293)*, 2(1), 32–43.
- Oliver, R. L. (1999). Whence Consumer Loyalty? *Journal of Marketing*, 63(4\_suppl1), 33–44. <https://doi.org/10.1177/00222429990634s105>
- Oliver, R. L. (2010). Customer satisfaction. *Wiley International Encyclopedia of Marketing*.
- Pangilinan, G. A., Tambunan, A., & Astuti, E. D. (2023). Tokopedia E-Commerce is Being Used to Present Opportunities for Young Business Owners to Succeed in the Digital Economy Amid the Pandemic. *Startupreneur Business Digital (SABDA Journal)*, 2(2), 182–191.
- Python, W. (2021). Python. *Python Releases for Windows*, 24.
- Ramdhan, M. (2021). *Metode penelitian*. Cipta Media Nusantara.
- Reichheld, F. F. (2003). The one number you need to grow. *Harvard Business Review*, 81(12), 46–55.
- Samsir, Irmayani, D., Edi, F., Harahap, J. M., Jupriaman, Rangkuti, R. K., Ulya, B., & Watrianthos, R. (2021). Naives Bayes Algorithm for Twitter Sentiment Analysis. *Journal of Physics: Conference Series*, 1933(1), 012019. <https://doi.org/10.1088/1742-6596/1933/1/012019>

- Saputri, Y. R., & Februariyanti, H. (2022). Sentiment analysis on shopee e-commerce using the naïve bayes classifier algorithm. *Jurnal Mantik*, 6(2), 1349–1357.
- Sugiharto, B. H. (2024). The role of e-commerce for MSMEs as a digital marketing strategy in facing industrial revolution 4.0. *Management Studies and Business Journal (PRODUCTIVITY)*, 1(1), 99–107.
- Thomas, D. M., & Mathur, S. (2019). Data analysis by web scraping using python. *2019 3rd International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, 450–454.
- Tri Jaka, A. H. (2015). *Preprocessing Text untuk Meminimalisir Kata yang Tidak Berarti dalam Proses Text Mining*.
- Viet, T. N., Le Minh, H., Hieu, L. C., & Anh, T. H. (2021a). The Naïve Bayes algorithm for learning data analytics. *Indian Journal of Computer Science and Engineering*, 12(4), 1038–1043.
- Viet, T. N., Le Minh, H., Hieu, L. C., & Anh, T. H. (2021b). The Naïve Bayes algorithm for learning data analytics. *Indian Journal of Computer Science and Engineering*, 12(4), 1038–1043.
- Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, 55(7), 5731–5780.
- Yusuf, M., Afifah Fitriyani, Z., Abdilah, A., Ardianto, R., Suhendar, A., & Pertiwi, U. (2022). *THE IMPACT OF USING TOKOPEDIA ON PROFITABILITY AND CONSUMER SERVICE: THE IMPACT OF USING TOKOPEDIA ON PROFITABILITY AND CONSUMER SERVICE* (Vol. 30, Issue 2).